T-STAT for Detecting Autism Spectrum Disorder in Toddlers Aged 18–24 Months

Abstract

Two studies were conducted to examine the use of the Taiwanese version of the Screening Tool for Autism in Two-Year-Olds (T-STAT) for detecting autism spectrum disorder (ASD) in toddlers aged 18–24 months. Study 1 used receiver operating characteristics to select a cutoff score for the T-STAT. It involved 2 groups of 16 toddlers each, 1 group with toddlers having ASD and the other group with chronological and mental age-matched toddlers with developmental delay (DD). The result suggested that a cutoff of 2.50 would yield high sensitivity and specificity. In Study 2, we recruited 136 toddlers—30 with ASD, 33 with Mild-ASD, and 73 with DD. Using 2.50 as the cutoff score, the concurrent agreement between T-STAT risk and clinical diagnosis and Autism Diagnostic Observation Schedule yielded high sensitivity and specificity. The results of this study indicated that the T-STAT can be used as a Level 2 autism-specific screening tool for the 18–24 months age group.

Keywords: autism spectrum disorder, toddler, screening, sensitivity, specificity

Autism spectrum disorder (ASD), a neurodevelopmental disorder, is characterized by impairments in social and communication skills, repetitive behavior patterns, and a restricted range of interests (American Psychiatric Association [APA], 2013). Studies have shown that early intervention can significantly improve impairments and long-term outcomes for children with ASD (Dawson et al., 2010; Pickles et al., 2016; Rogers et al., 2019). The effectiveness of early intervention highlights the importance of early detection and early diagnosis. ASD can be reliably diagnosed when a child is < 24 months of age (Barbaro & Dissanayake, 2017; Guthrie et al., 2013; Ozonoff et al., 2015; Zwaigenbaum et al., 2016). However, frequently, children with ASD are diagnosed after 3 years (Bent et al., 2015; Daniels & Mandell, 2014). Delayed diagnosis might equivalently lead to delayed access to early intervention and family support services (Bent et al., 2015). Children with ASD who received early intervention before the age of 3 years showed improved outcomes in targeted skills (e.g., social communication and imitation) (Zwaigenbaum et al., 2015). Delayed diagnosis and intervention can considerably affect the developmental outcomes of children with ASD. Thus, facilitating the early detection and early diagnosis of ASD is crucial, particularly in underresourced countries or communities.

Earlier, ASD was considered a rare disorder. However, the estimated prevalence of children with ASD has drastically increased by 1–2% (APA, 2013; Baio et al., 2018; Baron-Cohen et al., 2009; Christensen et al., 2016). ASD prevalence in Taiwan is lower than that in Western countries (Sun et al., 2013). Reasons for low ASD prevalence in Taiwan are as follows: (1) valid screening tools for toddlers are limited, and children with ASD are not detected in the clinical or community setting. (2) Parents do not continually bring their children with developmental problems to visit clinicians, and monitoring of signs of abnormalities is not conducted. (3) Consistent with Ward et al.'s (2016) study, many clinicians in Taiwan have little experience or knowledge regarding screening or diagnosing ASD in children < 3 years of age (Lai et al., 2012) and advice parents to "watch and wait".

Thus, it is crucial to develop an ASD screening tool that is cost-effective and easy to administer in toddlers for detecting early signs of ASD. This can facilitate early screening or early diagnosis of ASD in Taiwan.

Existing screening tools for ASD can be divided into two types, namely Levels 1 and 2 (Filipek et al., 1999; Fombonne, 2009). Level 1 screening tools are developed for use in the general population, whereas Level 2 screening tools are designed for use in individuals with a high ASD risk. In Taiwan, a majority of infants and toddlers regularly undergo physical and developmental surveillance in primary care settings (e.g., community clinics and health centers) during their vaccination. Routine checkups allow healthcare providers to examine infants and toddlers for socioemotional functioning (e.g., response to own name) and detect high-risk cases. Infants and toddlers with positive indicators of ASD are referred to specialty clinics (e.g., department of child psychiatry at regional hospitals) for a comprehensive assessment by a multidisciplinary team and for making a formal diagnosis. A Level 2 (rather than a Level 1) screening tool is needed in these clinical settings for differentiating between toddlers with ASD and those with other developmental problems. In Taiwan, medical certificates as ASD proof for children are mainly issued by child and adolescent psychiatrists. In 2010, a total of 210 licensed child and adolescent psychiatrists were available, and the ratio of child and adolescent psychiatrists to children was close to 1: 20,000 (National Health Research Institutes [NHRI], 2019). They typically shoulder a heavy workload due to a high number of patients per expert. Therefore, an affordable and easy-to-administer screening tool is essential to facilitate the early screening and early diagnosis of ASD among toddlers.

The Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001) is a parent-reported checklist that was originally developed as a Level 1 screening tool for ASD. This tool was translated and validated for use in Taiwan. The M-CHAT consists of 23 yes/no questions designed to evaluate the development and behavior of toddlers aged 16–30 months and can be completed by a parent/caregiver within 5–10 min. Wong et al. (2018) recruited

236 children, comprising 113 with ASD and 123 with developmental delay (DD), aged 18–47 months with a high ASD risk from a rural area of Southern Taiwan. Using "failing any 4 of the 23 M-CHAT items" as cutoff, it showed a sensitivity of 0.77 and a specificity of 0.72. According to the standard suggested by Cicchetti et al. (1995), values < 0.70, between 0.70 and 0.79, between 0.80 and 0.89, and > 0.90 indicate poor, fair, good, and excellent accuracy, respectively. The diagnostic accuracy of the M-CHAT was thus fair in this high-risk sample. This result can be attributed to factors such as the stigma that is associated with psychological diagnoses in Chinese culture (Pang et al., 2018; Wong et al., 2018) and/or the lack of sufficient knowledge about ASD in parents (Wong et al., 2018). If this is the case, using an interactive screening tool that provides clinicians an opportunity to directly observe behaviors of infants and toddlers could improve screening accuracy.

Contrary to the M-CHAT used to detect ASD in the general population, the Screening Tool for Autism in Two-Year-Olds (STAT; Stone et al., 2004) is an interactive Level 2 screening tool for detecting ASD in high-risk individuals. Previous studies (e.g., Stone et al., 2008; Wu & Chiang, 2014) have suggested that toddlers with ASD show social-communicative impairments. Thus, social-communicative deficits could be indicators used for early screening of ASD. The STAT consists of 12 activity-based items that measure four social-communicative domains: play, requesting, directing attention (i.e., joint attention), and imitation. It is a brief screening tool used by professionals and is likely to be useful in clinical settings. The STAT was originally designed for use among children aged 24–35 months. Stone et al. (2004) recruited 104 young children with developmental problems, consisting of 65 young children with ASD and 39 young children with DD. Using a cutoff score of 2, the STAT yielded excellent sensitivity (1) and specificity (0.90). Stone et al. (2008) further tested the STAT's usage for ASD screening among toddlers aged < 24 months. They recruited 71 toddlers aged 12–23 months, consisting of 19 toddlers with ASD and 52 toddlers without ASD (i.e., DD, language impairment, broad autism phenotype, and no concerns). A

cutoff score of 2.75 on the STAT yielded excellent sensitivity (0.95) and fair specificity (0.73). Their findings suggested that ASD could be detected in toddlers at < 24 months of age. However, their findings suggested a high number of false positives in toddlers aged < 13 months. Recently, Wu et al. (2020) examined the use of the STAT for detecting ASD in toddlers aged < 24 months. They recruited 119 toddlers (57 with ASD and 62 with DD) aged 16–24 months (Time 1) and finalized diagnosis at 18 months after Time 1. A cutoff score of 2.5 on the STAT yielded good prospective sensitivity (0.86) and specificity (0.81). The difference in the age range of participants could be one possible reason for different cutoffs in these two studies.

The Taiwanese version of the STAT was developed and called the Taiwanese version of the Screening Tool for Autism in Two-Year-Olds (T-STAT; Chiang et al., 2013), which also consists of 12 activity-based items that measure four social-communicative domains: play, requesting, joint attention, and imitation. Chiang and colleagues (2013) recruited 43 young children with ASD and 34 young children with DD. With a cutoff score of 2, the T-STAT yielded good sensitivity (0.86) and specificity (0.82). Wu and colleagues (2019) further tested the T-STAT's usage in ASD screening in children aged 36–48 months. They recruited 84 and 63 children with ASD and DD, respectively. With a cutoff score of 1.25, the T-STAT yielded good sensitivity (0.89) and excellent specificity (0.92). The T-STAT is therefore a promising Level 2 screening tool for ASD in young children. These initial studies demonstrated that the T-STAT is a promising Level 2 screening tool for ASD in children aged 36 months and older.

To extend the T-STAT's use in early detection and diagnosis, we conducted two studies that investigated whether it could be used for detecting ASD in toddlers aged 18–24 months by (1) using receiver operating characteristics (ROC) to compute the potential cutoff of the T-STAT in the development sample, and by (2) examining the cutoff and validity of the T-STAT in the validation sample.

Study 1: Computing Cutoff Scores of the T-STAT

Methods

Participants

This study was approved by the Ditmanson Medical Foundation Chia-Yi Christian Hospital Research Ethics Committee. All parents provided informed consent before the assessment. In total, 32 toddlers, 16 with ASD and 16 with DD, aged 18–24 months participated in the study. Of the 32 participants, 22 were a part of the participant sample in the study by Wu et al. (2020). None of the participants had sensory or motor impairments or previously diagnosed genetic disorders. All participants were diagnosed with either ASD or DD based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fifth edition (DSM-5; APA, 2013). According to the DSM-5 criteria for ASD, a child must exhibit a minimum of three deficits in social-communicative/interaction skills and two restricted/repetitive behaviors. All participants with ASD were assessed and diagnosed by a multidisciplinary team consisting of two senior clinical child psychologists with doctoral degrees and two senior child and adolescent psychiatrists. Their diagnoses were based on participant's developmental history, parental concerns, cognitive and adaptive functioning, clinical observations, and results of Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999). The participants were considered to have DD if they did not meet the DSM-5 criteria for ASD and failed to reach a total score of 85 on the Mullen Scales of Early Learning (MSEL; Mullen, 1995) or had a *T*-score of 35 on any of the four cognitive scales (i.e., visual reception, fine motor, receptive language, and expressive language).

Mental ages (MAs) of all participants were computed by averaging the age equivalents across the four cognitive scales of the MSEL (Mullen, 1995). Independent-samples *t* tests showed comparable MAs, chronological age, and parents' years of education in the ASD and DD groups. Furthermore, the two groups did not differ in their sex ratios. Nevertheless,

toddlers with ASD obtained higher scores on the ADOS than toddlers with DD. The demographic characteristics of the sample are presented in Table 1.

Insert Table 1 about here

Procedures and Measures

All participants were assessed using the T-STAT (Chiang et al., 2013; Wu et al., 2019), MSEL (Mullen, 1995), and ADOS (Lord et al., 1999). The test administration conducted in the current study was the same as that in Wu et al. (2020). That is, a total of 14 items were included in both the STAT and T-STAT. The only difference between these two studies is the scoring algorithm of the data used. The T-STAT was administered by examiners who were graduate students in the Master of Science degree program in the discipline of clinical psychology and had received 8-hr training for administering and scoring the assessment. They were blinded to the diagnostic information of the participants as well as to the concerns of the caregivers before administration. Inter-rater reliability between these examiners and the first author who was trained for administering and scoring the T-STAT was high (i.e., 0.90). ADOS was administered by the authors who had received research training and certification in Taiwan (i.e., by Dr. Catherine Rice's team at Pingtung county). Furthermore, they were not provided with any information regarding the T-STAT before the administration procedure. All examiners periodically discussed the manner in which they scored the T-STAT and ADOS to further ensure inter-rater reliability.

Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a standardized developmental test that was designed for preschool children aged 0–68 months. It consists of four cognitive scales: visual reception, fine motor, receptive language, and expressive language. The four cognitive scales yield *T*-scores, which have a mean of 50. The four

subscale scores can be used to compute a composite score, which is an indicator of early learning and has a mean of 100. The MSEL has demonstrated concurrent validity against other well-known developmental tests of language and cognitive development (e.g., Bayley Scales of Infant Development; Bayley, 1969). In addition, it has demonstrated acceptable internal consistency and test–retest reliability.

Taiwanese Version of Screening Tool for Autism in Two-Year-Olds (T-STAT; Chiang et al., 2013; Wu et al., 2019). The T-STAT is an interactive measure that was originally designed to screen for autism in children aged 24–35 months. The T-STAT is an individually administered assessment that consists of 12 activity-based items, and it takes approximately 20 min to complete. It measures four early social-communicative skills: play (two items), requesting (two items), joint attention (four items), and imitation (four items). All of the items are scored as either pass or fail. The number of failure items in each domain is converted into scores. The scores for the two-item domains can be 0, 0.5, or 1, whereas the scores for the four-item domains can be 0, 0.25, 0.50, 0.75, or 1. Thus, the scores for each domain of the T-STAT can range from 0 to 1. In addition, the total T-STAT score can be computed by summing of the four domain scores. Therefore, the composite score can range from 0 to 4; higher scores are indicative of greater impairment levels. The T-STAT has demonstrated a good level of accuracy in identifying ASD and DD in children aged 24–48 months (Chiang et al., 2013; Wu et al., 2019).

Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999). The ADOS is a semi-structured play-based and observational assessment, which is divided into four modules. Each module is selected based on the age and expressive language of the respondent. The ADOS is considered the best diagnostic tool for ASD because it serves as a standardized

means of observing and scoring language and communication skills, reciprocal social and stereotypic behaviors, and restricted interests. Each module provides an algorithm that entails cutoffs that can be used to assign respondents to one of the following three categories: autism, ASD (i.e., pervasive developmental disorder-not otherwise specified; PDD-NOS), or non-ASD. In the present study, both autism and PDD-NOS were merged into one category, namely ASD. Because this study had relatively young participants, only Module 1 was administered. A modified version of the ADOS for toddlers aged 12–30 months, namely the ADOS-2: Toddler Module (Lord et al., 2012), was not used in the present study because it is yet to be culturally adapted and validated for Taiwan.

Analysis

Statistical Package for the Social Sciences was used to conduct statistical analyses in this study. The screening properties of the T-STAT were examined using ROC. ROC was examined to select the optimal range of cutoff scores of the T-STAT and consequently to examine sensitivity, specificity, and the area under the curve (AUC).

Results

The results of an ROC suggested that the optimal cutoff scores for better sensitivity and specificity were between 2 and 3. Sensitivity and specificity associated with different cutoff scores for the sample are presented in Table 2. A score of ≥ 2.50 was then selected as the cutoff for ASD risk.

With 2.50 as the cutoff score for ASD risk, the results derived from sensitivity, specificity, the positive predictive value, and the negative predictive value were all the same (100%) (Table 3). Both sensitivity and specificity indicated the excellent validity of the T-STAT, as suggested by Cicchetti and colleagues (1995). ROC yielded an AUC of 1, which also demonstrated excellent classification accuracy. The initial findings supported that the T-STAT could be used to detect ASD in toddlers aged 18–24 months.

In addition, independent *t* tests were used to compare the T-STAT total scores between the two groups. The results showed that toddlers with ASD (mean = 3.41, standard deviation [SD] = 0.45) obtained a significantly higher score than toddlers with DD (mean = 1.56, SD = 0.41), t (30) = 12.13, $p < 0.001$).

Insert Table 2 about here

Insert Table 3 about here

STUDY 2: Validity of the T-STAT in the Validation Sample

In Study 1, a T-STAT score of 2.50 was found to be the cutoff for ASD risk based on ROC results. In Study 2, we recruited more participants independent of the Study 1 sample to test the cutoff criteria and classification accuracy.

Method

Participants

Participants were recruited from the southwest area of Taiwan. They were diagnosed based on a different set of criteria than those used in Study 1. This was because studies (e.g., Frazier et al., 2012) have shown that the criteria of DSM-5 have a lower sensitivity than those of Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (DSM-IV-TR; APA, 2000). In other words, participants showing significant impairments related to the core ASD symptoms but failing to meet the full criteria for ASD according to DSM-5 are misclassified as having DD, which is not appropriate. Therefore, Frazier et al. (2012) proposed a set of less stringent criteria than the DSM-5 criteria for ASD. Accordingly, the following criteria were used in Study 2: (1) three deficits in social-communicative/interaction skills and one restricted/repetitive behavior and (2) two deficits in social-communicative/interaction skills and two restricted/repetitive behaviors.

Toddlers who met these less stringent DSM-5 criteria were classified as having Mild-ASD. Finally, 30, 33, and 73 toddlers with ASD, Mild-ASD, and DD, respectively, participated. Of the 136 participants, 89 had participated in the study by Wu et al. (2020).

Table 4 presents the demographic characteristics of participants in Study 2. One-way analysis of variance was used to determine whether the three groups differed significantly in terms of demographic characteristics. The results revealed no significant difference in parents' years of education and sex ratio.

Insert Table 4 about here

Procedures and Measures

All toddlers received an assessment battery that included the MSEL, T-STAT, and ADOS. An informed consent form was signed by parents before the measures were administered. As in Study 1, the T-STAT was administered by graduate students of clinical psychology in the Master of Science degree program who had received training for administering and scoring in advance. Furthermore, the T-STAT and ADOS were administrated by different examiners who were blinded to each other's testing results. The clinical diagnosis was made by senior psychiatrists or psychologists who were trained, and they achieved excellent inter-rater reliability based on interviews of caregivers and observations made throughout the evaluation. Moreover, they were blinded to participants' T-STAT risk status. The cutoff score for the T-STAT based on the results of Study 1 was used in Study 2.

Results

The three groups' performance on the T-STAT is shown in Table 5. Given that the three groups were unmatched by MAs, analysis of covariance was further executed for testing

performance differences of the three groups on the T-STAT. Using MAs as a covariate based on clinical diagnosis, the results revealed that significant group differences existed for scores of the play, requesting, and joint attention domains and total score of the T-STAT (Table 6).

Insert Table 5 about here

Insert Table 6 about here

Concurrent validity of the T-STAT was examined by comparing the children's T-STAT risk category with their clinical diagnosis and ADOS classification (Table 7). Using a cutoff score of 2.50 for high risk, accurate diagnoses of ASD or Mild-ASD in 58 of 63 toddlers and DD in 63 of 73 toddlers were made in comparison with their clinical diagnoses. Sensitivity was 0.92 and specificity was 0.86 for the T-STAT classification of ASD (or Mild-ASD) and DD. Five toddlers with Mild-ASD and 10 toddlers with DD were misidentified using the T-STAT. Both ADOS and T-STAT provide categorical classification of risk. Thus, this study also compared the agreement between categories of ADOS and T-STAT (Table 7). The results showed that 66 of 74 toddlers and 60 of 62 toddlers were accurately diagnosed with autism (or PDD-NOS) and DD, respectively, compared with their ADOS classification. Sensitivity was 0.89 and specificity was 0.97 for the T-STAT classification of autism (or PDD-NOS) and DD. Three toddlers with autism, five toddlers with PDD-NOS, and two toddlers with DD were misidentified using the T-STAT.

Insert Table 7 about here

According to clinical diagnosis, ROC yielded AUCs of 0.95 (confidence interval = 0.91–0.99) and 0.90 (confidence interval = 0.84–0.95) for ASD versus DD and ASD + Mild-ASD versus DD, respectively. According to ADOS classification, ROC yielded AUCs

of 0.95 (confidence interval = 0.90–1) and 0.92 (confidence interval = 0.86–0.97) for autism versus DD and autism + PDD-NOS versus DD, respectively. The findings suggested that the T-STAT had good screening accuracy and could be used as a screening tool for detecting ASD in toddlers aged < 24 months in the clinical setting in Taiwan.

Discussion

Level 2 screening tools for detecting ASD in toddlers < 24 months in high-risk samples in Taiwan are few. Parent-reported screening tools (e.g., M-CHAT) are advantageous, as they are cost-effective and convenient. However, a few factors (e.g., parents' misunderstanding of the questions and stigma) may lead to low accuracy, especially for underresourced communities. Interactive screening tools such as the T-STAT developed by Chiang et al. (2013) could be useful for detecting autism in toddlers aged 24–35 months. Thus, the purpose of the present study was to expand the utility of the T-STAT and enhance early screening for ASD in a large age range according to the needs of current clinical services in Taiwan. Two studies were conducted to examine the validity of the T-STAT for identifying ASD in toddlers aged 18–24 months.

In accordance with past findings (e.g., Barbaro & Dissanayake, 2017; Guthrie et al., 2013), toddlers with ASD can be early screening and early diagnosis at age < 24 months. Studies have shown that cutoff scores of 2 and 1.25 on the T-STAT can be used to reliably identify autism in children aged 24–35 months (Chiang et al., 2013) and 36–48 months (Wu et al., 2019), respectively. Compared with previous studies (i.e., Chiang et al., 2013; Wu et al., 2019), our study had younger participants. The results of the current study showed that a high cutoff score of 2.50 on the T-STAT could exhibit good or excellent sensitivity (0.89–1) and specificity (0.86–0.97) in differentiating toddlers aged 18–24 months with ASD/Mild-ASD from those with DD. In addition, an examination of the AUC (0.90–0.95) revealed that the T-STAT can reliably identify ASD/Mild-ASD in toddlers aged 18–24 months. The results of

this study suggest that for accurate diagnosis of ASD, cutoff scores of 2.50 on the T-STAT must be used for toddlers aged 18–24 months.

The current study and Wu et al. (2020) found that using a cutoff score of 2.50, instead of 2.75 as recommended by Stone et al. (2008), would yield better discrimination. Stone and colleagues proposed that the younger the children with ASD are, the more the deficits of early social-communicative skills. They recruited a sample of toddlers aged 12–23 months, whereas our sample and Wu et al.' sample were slightly older (i.e., 16–24 months). Thus, it is reasonable for the cutoff score to be lower than that of Stone et al. This argument was supported by the findings of the current study and those of Chiang and colleagues (2013) and Wu and colleagues (2019), all of which suggested significant age-related developments in early social-communicative skills across preschool children with ASD. It is necessary to consider age-related developments when using the severity level of early social-communicative skills to detect ASD in children.

Similar to previous studies (e.g., Stone et al., 2008; Veness et al., 2012), in this study, toddlers with ASD demonstrated deficits in early social-communicative skills at < 24 months of age. Even after controlling for MAs, early social-communicative impairments were evident in toddlers with ASD. Similar to a study by Wu and colleagues (2019), our study showed that toddlers with ASD showed deficits in the joint attention domain, followed by the requesting, play and imitation domains. The findings of this study suggested that integrations of multiple nonverbal communication skills (e.g., coordinated eye contact and gesture/vocalization) can be used to differentiate toddlers aged < 24 months with ASD from those with DD. In accordance with previous studies (e.g., Wu et al., 2019), the findings of this study also showed that the imitation domain is a weak discriminator for distinguishing toddlers with ASD from those with DD. This is consistent with previous findings that children with ASD do not exhibit significant impairments in tasks that require the imitation of meaningful actions that involve objects (e.g., Hepburn & Stone, 2006; Stone et al., 1997; Wu

& Chiang, 2014). Given that the imitation domain has two items that necessitate imitation of meaningful actions that involve objects, it is not a robust discriminator. In future, it might be beneficial to replace the insensitive items in the imitation domain.

Frazier and colleagues (2012) suggested that the DSM-5 criteria have a lower sensitivity than the criteria in the former edition. Individuals who did not meet the DSM-5 criteria for ASD may still have significant impairments related to the core ASD symptoms. Thus, one strength of this study relative to previous T-STAT studies is the inclusion of toddlers with ASD or Mild-ASD using the strict and relaxed DSM-5 criteria for ASD simultaneously. Our findings suggested that toddlers with Mild-ASD showed milder autism symptomatology than those with ASD. Toddlers with Mild-ASD had higher scores and exceeded cutoff scores on both the T-STAT and ADOS; thus, they might be regarded as having ASD instead of DD. However, these toddlers must be followed up for confirming their diagnosis.

Mild-ASD, which was not included in Study 1, had lower total scores than those with ASD in Study 2. Thus, the cutoff score was decreased, and accuracy was examined again. When using 1.75, 2, and 2.25 as cutoff for the T-STAT, respectively, 37 (51%), 26 (36%), and 13 (18%) individuals with DD and 3 (9%), 4 (12%), and 5 (15%) individuals with Mild-ASD were misidentified. The results of this study indicated that cutoff scores might need to be lowered (e.g., 2) for detecting Mild-ASD. Furthermore, clinicians must collect other information or execute a comprehensive assessment for early diagnosis of Mild-ASD because toddlers with DD tend to be misidentified.

In this study, 10 toddlers with DD were misidentified as having ASD using the T-STAT. Among these toddlers, eight met the ADOS criteria for ASD. In addition, one toddler with a clinical diagnosis of DD was identified as having a high risk of ASD on the T-STAT due to his shyness and anxiety during the assessment. The findings suggested that child characteristics (e.g., shyness and anxiety) might have caused the higher rate of item failures

on the T-STAT. Using a high cutoff could increase specificity and reduce false positive results. When using a high score (i.e., 2.75) as cutoff for the T-STAT, specificity increased from 0.86 to 0.88, whereas sensitivity decreased from 1 to 0.77 and from 0.85 to 0.76 for ASD and Mild-ASD, respectively. Compared with false positive results, false negative results may lead to costly outcomes for toddlers with ASD, their families, and society (Stone et al., 2008), such as misunderstanding children's behaviors and missing out from early intervention. Thus, using a cutoff score of 2.50 on the T-STAT is acceptable despite slightly low specificity. If clinicians are still concerned regarding false positive cases, the T-STAT can be combined with an interview or parent-reported screening tools to enhance diagnostic accuracy. Diagnosis could not be confirmed in only a minority of referred children, and professionals able to conduct such assessments are limited in most communities (Zwaigenbaum & Warren, 2020). Zwaigenbaum and Warren suggested that the limited reserve of expert and comprehensive assessment may be efficiently used to serve these children with ambiguous diagnosis. In most communities, including Taiwan, professionals are limited and shoulder a heavy workload. Healthcare professionals can be trained for using the T-STAT for detecting ASD in toddlers. Then, toddlers without clear autistic symptoms (e.g., total scores of the T-STAT are 2.25 or 2) can be referred for a comprehensive evaluation and continual monitoring. For toddlers with clear autistic symptoms (e.g., total scores of the T-STAT are 2.75 or 3), professionals can diagnose them early and provide early intervention.

The present study investigated the use of the T-STAT in Taiwan among at-risk toddlers aged < 24 months. The study findings suggested that the T-STAT can reliably detect at-risk children with ASD within the developmental period ranging from toddlerhood to preschool age. The T-STAT is a Level 2 interactive screening tool that can be completed within 20 min and is easy to administer. It may be promoted among practitioners (e.g., clinical

psychologists and occupation therapists) in clinical settings (e.g., regional hospitals) to differentiate toddlers with ASD from those with DD for making formal ASD diagnosis. Then, evidence-based interventions can be provided for toddlers with ASD and their families.

Limitations and Future Directions

In conclusion, the present study used the T-STAT as a Level 2 screener for ASD among at-risk toddlers aged 18–24 months. The results suggest that the T-STAT has a good or excellent level of concurrent validity (e.g., sensitivity and specificity) and can therefore be used as an autism-specific screening tool for children ranging from toddlerhood to preschool age. However, this study has a few limitations. First, given that professionals have a heavy workload, it needs to develop a brief version of the T-STAT for detecting ASD in infants and children in Taiwan. It can be helpful for early screening and early diagnosis. Second, the current study had a cross-sectional design. Longitudinal studies are needed for examining change in and stability of the T-STAT risk category, especially for long-term follow-up (e.g., 5 years). Third, contrary to previous studies (e.g., Stone et al., 2008), this study did not include toddlers aged 14–17 months. Thus, the findings only supported that the T-STAT could be used to detect ASD in toddlers aged 18–24 months. For early screening and early diagnosis of ASD in the young population, future research is needed and should include toddlers with ASD and those with DD aged 14–17 months to further examine and validate the cutoffs of the T-STAT. Fourth, this study was executed in a rural agricultural area of Southern Taiwan. Therefore, similar to previous studies (Chiang et al., 2013; Wu et al., 2019), our study encountered difficulties in recruiting a sample with an ideal size. Hence, in Study 1, our sample could not be used to examine the scoring algorithm and to investigate its validity simultaneously. In addition, Mild-ASD was not included as a distinct group for deciding cutoffs. Accordingly, recruiting participants from the urban areas of Taiwan and validating the T-STAT with a large sample are strongly recommended in future research.

Compliance with Ethical Standards

Funding: This study was supported by the Ministry of Science and Technology (formerly National Science Council) (MOST-103-2628-H-037-001-MY2; MOST-105-2410-H-037-001-MY3).

Ethical Approval: In the present study, all the procedures that involved human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee as well as the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the Ditmanson Medical Foundation Chia-Yi Christian Hospital Research Ethics Committee (CYCH-IRB101022; CYCH-IRB102045).

Informed Consent: Informed consent was obtained from all the individuals who participated in this study.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). American Psychiatric Publishing.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Bayley, N. (1969). *Bayley Scales of Infant Development*. The Psychological Corporation.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L.-C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., … Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years−Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR Surveillance Summaries, 67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Barbaro, J., & Dissanayake, C. (2017). Diagnostic stability of autism spectrum disorder in toddlers prospectively identified in a community-based setting: Behavioural characteristics and predictors of change over time. *Autism, 21*(7), 830–840. https://doi.org/10.1177/1362361316654084

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry, 194*(6), 500–509. https://doi.org/10.1192/bjp.bp.108.059345

Bent, C. A., Dissanayake, C., & Barbaro, J. (2015). Mapping the diagnosis of autism spectrum disorders in children age under 7 years in Australia, 2010–2012. *The Medical Journal of Australia, 202*(6), 317–320. https://doi.org/10.5694/mja14.00328

Formatted: French (France)

Chiang, C.-H., Wu, C.-C., Hou, Y.-M., Chu, C.-L., Liu, J.-H., & Soong, W.-T. (2013). Development of T-STAT for early autism screening. *Journal of Autism and Developmental Disorders, 43*(5), 1028–1037. https://doi.org/10.1007/s10803-012-1643-4

Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkin, M. S., Fitzgerald, R. T., Rice, C., Kurzius-Spencer, M., Baio, J., & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *Journal of Developmental and Behavioral Pediatrics*, *37*(1), 1–8. https://doi.org/10.1097/DBP.0000000000000235

Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychology, 1*(1), 26–37. https://doi.org/10.1080/09297049508401340

Daniels, A. M., & Mandell, D. S. (2014). Explaining differences in age at autism spectrum disorder diagnosis: A critical review. *Autism, 18*(5), 583–597. https://doi.org/10.1177/1362361313480277

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver Model. *Pediatrics, 125*(1), e17–e23. https://doi.org/10.1542/peds.2009-0958

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Jr., Dawson, G., Gordon, B., Gravel, J. S., Johnson, C. P., Kallen, R. J., Levy, S. E., Minshew, N. J., Prizant, B. M., Rapin, I., Rogers, S. J., Stone, W. L., Teplin, S., Tuchman, R. F., & Volkmar, F. R. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*(6), 439–484. https://doi.org/10.1023/A:1021943802493

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research, 65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., Constantino, J., Findling, R. L., Hardan, A. Y., & Eng, C. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(1), 28–40.e3. https://doi.org/10.1016/j.jaac.2011.09.021

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry, 54*(5), 582–590. https://doi.org/10.1111/jcpp.12008

Hepburn, S. L., & Stone, W. L. (2006). Longitudinal research on motor imitation in autism. In S. J. Rogers & J. H. G. Williams (Eds.), *Imitation and the social mind: Autism and typical development* (pp. 310–328). The Guilford Press.

Lai, D.-C., Tseng, Y.-C., Hou, Y.-M., & Guo, H.-R. (2012). Gender and geographic differences in the prevalence of autism spectrum disorders in children: Analysis of data from the national disability registry of Taiwan. *Research in Developmental Disabilities*, *33*(3), 909–915. https://doi.org/10.1016/j.ridd.2011.12.015

Lord, C., Luyster, R. J., Gotham, K., & Guthrie, W. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part II): Toddlers Module*. Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Western Psychological Services.

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS Ed.). American Guidance Service.

National Health Research Institutes (2019). *A proposal on children's medical care and health policy.* National Health Research Institutes. https://chrc.nhri.org.tw/professionals/achieve.html

Formatted: German (Germany)

Formatted: English (United States)

Ozonoff, S., Young, G. S., Landa, R. J., Brian, J., Bryson, S., Charman, T., Chawarska, K., Macari, S. L., Messinger, D., Stone, W. L., Zwaigenbaum, L., & Iosif, A.-M. (2015). Diagnostic stability in young children at risk for autism spectrum disorder: A baby siblings research consortium study. *Journal of Child Psychology and Psychiatry, 56*(9), 988–998. https://doi.org/10.1111/jcpp.12421

Pang, Y. L., Lee, C. M., Wright, M., Shen, J., Shen, B., & Bo, J. (2018). Challenges of case identification and diagnosis of autism spectrum disorders in China: A critical review of procedures, assessment, and diagnostic criteria. *Research in Autism Spectrum Disorders, 53*, 53–66. https://doi.org/10.1016/j.rasd.2018.06.003

Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., Gammer, I., Lowry. J., Vamvakas. G., Byford. S., Aldred. C., Slonims, V., McConachie, H., Howlin, P., Parr, J. R., Charman, T., & Green, J. (2016). Parent-mediated social communication therapy for young children with autism (PACT): Long-term follow-up of a randomised controlled trial. *The Lancet, 388*(10059), 2501–2509. https://doi.org/10.1016/S0140-6736(16)31229-6

Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*(2), 131–144. https://doi.org/10.1023/a:1010738829569

Rogers, S. J., Estes, A., Vismara, L., Munson. J., Zierhut.C., Greenson, J., Dawson, G., Rocha, M., Sugar, C., Senturk, D., Whelan, F., & Talbott, M. (2019). Enhancing low-intensity coaching in parent implemented Early Start Denver Model intervention for early autism: A randomized comparison treatment trial. *Journal of Autism and Developmental Disorders, 49*(2), 632–646. https://doi.org/10.1007/s10803-018-3740-5

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental*

Formatted: French (France)

Disorders, *34*(6), 691–701. https://doi.org/10.1007/s10803-004-5289-8

Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in two-year-olds (STAT) for children under 24 months: An exploratory study. *Autism*, *12*(5), 557–573. https://doi.org/10.1177/1362361308096403

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology, 25*(6), 475–485. https://doi.org/10.1023/a:1022685731726

Sun, X., Allison, C., Matthews, F. E., Sharp, S. J., Auyeung, B., Baron-Cohen, S., & Brayne, C. (2013). Prevalence of autism in mainland China, Hong Kong and Taiwan: A systematic review and meta-analysis. *Molecular Autism, 4*, 7. https://doi.org/10.1186/2040-2392-4-7

Veness, C., Prior, M., Bavin, E., Eadie, P., Cini, E., & Reilly, S. (2012). Early indicators of autism spectrum disorders at 12 and 24 months of age: A prospective, longitudinal comparative study. *Autism, 16*(2), 163–177. https://doi.org/10.1177/1362361311399936

Ward, S. L., Sullivan, K. A., & Gilmore, L. (2016). Practitioner perceptions of the assessment and diagnosis of autism in Australia. *Australian Psychologist, 51*(4), 272–279. https://doi.org/10.1111/ap.12211

Wong, Y. S., Yang, C.-C., Stewart, L., Chiang, C.-H., Wu, C.-C., & Iao, L.-S. (2018). Use of the Chinese version modified checklist for autism in toddlers in a high-risk sample in Taiwan. *Research in Autism Spectrum Disorders*, *49*, 56–64. https://doi.org/10.1016/j.rasd.2018.01.010

Wu, C.-C., & Chiang, C.-H. (2014). The developmental sequence of social-communicative skills in young children with autism: A longitudinal study. *Autism, 18*(4), 385–392. https://doi.org/10.1177/1362361313479832

Wu, C.-C., Chiang, C.-H., Hou, Y.-M., Chu, C.-L., & Liu, J.-H. (2019). Utility of the Taiwan version of the screening tool for autism in two-year-olds to detect autism in children

Formatted: Italian (Italy)

Formatted: German (Germany)

Formatted: German (Germany)

Formatted: German (Germany)

Field Code Changed

aged three years. *Journal of Intellectual and Developmental Disability, 44*(3)*,* 337–345. https://doi.org/10.3109/13668250.2017.1413078

Wu, C.-C., Chu, C.-L., Stewart, L., Chiang, C.-H., Hou, Y.-M., & Liu, J.-H. (2020). The utility of the screening tool for autism in two-year-olds in detecting autism in Taiwanese toddlers who are less than 24 months of age: A longitudinal study. *Journal of Autism and Developmental Disorders, 50*, 1172–1181. https://doi.org/10.1007/s10803-019-04350-0

Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Fein, D., Kasari, C., Pierce, K., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., McPartland, J. C., Natowicz, M. R., Buie, T., Carter, A., Davis, P. A., Granpeesheh, D., Mailloux, Z., Newschaffer, C., Robins, D., … Wetherby, A. (2015). Early identification and interventions for autism spectrum disorder: Executive summary. *Pediatrics, 136*(Supplement 1), S1–S9. http://doi.org/10.1542/peds.2014-3667B

Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., Szatmari, P., Roncadin, C., Garon, N., & Vaillancourt, T. (2016). Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. *Autism Research, 9*(7), 790–800. https://doi.org/10.1002/aur.1585

Zwaigenbaum, L., & Warren, Z. (2020). Commentary: Embracing innovation is necessary to improve assessment and care for individuals with ASD: A reflection on Kanne and Bishop (2020). *The Journal of Child Psychology and Psychiatry*. Advance online publication. https://doi.org/10.1111/jcpp.13271

Table 1

Demographic Characteristics of the Participants

Variable	ASD (n = 16)	DD (n = 16)	p
CA (months)			
Mean (SD)	20.19 (1.28)	19.94 (1.65)	0.635
MAs (months)			
Mean (SD)	13.50 (1.53)	13.63 (0.89)	0.780
Parents' years of education			
Mean (SD): mother	14.75 (1.91)	13.31 (2.89)	0.108
Mean (SD): father	14.69 (3.20)	13.00 (2.81)	0.123
ADOS total scores			
Mean (SD)	18.50 (2.03)	3.94 (1.77)	< 0.001
Gender Ratio			
Male: Female	14:2	14:2	1

Note. CA= chronological age; MAs = mental ages; ADOS = Autism Diagnostic Observation Schedule; ASD = autism spectrum disorder; DD = developmental delay.

Table 2

Sensitivity and Specificity of Different T-STAT Cutoff Scores

Cutoff[a]	Sensitivity	Specificity
1.50	1.00	0.31
1.75	1.00	0.56
2.00	1.00	0.81
2.25	1.00	0.88
2.50	1.00	1.00
2.75	0.94	1.00
3.00	0.88	1.00

Note. [a]A score that is greater than or equal to the cutoff score indicates a risk of autism spectrum disorder.

Table 3

Classification Comparison Between T-STAT and Clinical Diagnosis

T-STAT risk category	ASD (n = 16)	DD (n = 16)
High risk	16 (100%)	0 (0%)
Low risk	0 (0%)	16 (100%)

Table 4

Demographic Characteristics of the Three Groups

	ASD (n = 30)	Mild-ASD (n = 33)	DD (n = 73)	p	Group difference
CA (months)					
Mean (SD)	22.67 (1.63)	21.58 (1.75)	21.22 (1.88)	0.001	ASD, Mild-ASD > DD
MAs (months)					
Mean (SD)	12.99 (3.03)	14.92 (3.30)	17.23 (3.01)	< 0.001	ASD, Mild-ASD < DD
Parents' years of education					
Mean (SD): mother	13.80 (2.83)	14.21 (2.53)	13.93 (2.60)	0.812	
Mean (SD): father	14.00 (2.45)	13.91 (2.97)	13.86 (2.40)	0.970	
ADOS total scores					
Mean (SD)	17.77 (3.21)	14.97 (3.51)	4.90 (3.57)	< 0.001	ASD > Mild-ASD > DD
Gender					
Male: female	28:2	26:7	50:23	0.024	

Note. CA= chronological age; MAs = mental ages; ADOS = Autism Diagnostic Observation Schedule; ASD = autism spectrum disorder; DD = developmental delay.

Table 5

Performance of T-STAT in the Three Groups

	ASD (n = 30)	Mild-ASD (n = 33)	DD (n = 73)	p	Group difference	Effect size
Play						
Mean (SD)	0.85 (0.30)	0.71 (0.31)	0.38 (0.35)	<0.001	ASD, Mild-ASD > DD	0.279
Requesting						
Mean (SD)	0.88 (0.28)	0.76 (0.36)	0.29 (0.37)	< 0.001	ASD, Mild-ASD > DD	0.373
Joint Attention						
Mean (SD)	0.88 (0.18)	0.72 (0.23)	0.34 (0.27)	< 0.001	ASD > Mild-ASD > DD	0.488
Imitation						
Mean (SD)	0.85 (0.17)	0.76 (0.20)	0.66 (0.24)	< 0.001	ASD > DD	0.115
Total Score						
Mean (SD)	3.47 (0.59)	2.95 (0.81)	1.67 (0.74)	< 0.001	ASD > Mild-ASD > DD	0.539

Table 6

Adjusted Performance of T-STAT in the Three Groups[1]

	ASD (n = 30)	Mild-ASD (n = 33)	DD (n = 73)	p	Group difference	Effect size
Play						
Mean (SD)	0.73 (0.06)	0.68 (0.05)	0.45 (0.04)	< 0.001	ASD, Mild-ASD > DD	0.127
Requesting						
Mean (SD)	0.83 (0.07)	0.74 (0.06)	0.32 (0.04)	< 0.001	ASD, Mild-ASD > DD	0.270
Joint Attention						
Mean (SD)	0.79 (0.05)	0.70 (0.04)	0.38 (0.03)	< 0.001	ASD, Mild-ASD > DD	0.343
Imitation						
Mean (SD)	0.78 (0.04)	0.74 (0.04)	0.70 (0.03)	0.223		0.019
Total Score						
Mean (SD)	3.15 (0.13)	2.85 (0.11)	1.85 (0.08)	< 0.001	ASD, Mild-ASD > DD	0.397

Note. [1]Adjusted for mental ages (MAs)

Table 7

Concurrent Validity of the T-STAT Category with Clinical Diagnosis and Autism Diagnostic Observation Schedule (ADOS) Classification

	Clinical diagnosis		
	ASD	Mild-ASD	DD
T-STAT risk category	(*n* = 30)	(*n* = 33)	(*n* = 73)
High risk	30 (100%)	28 (84.8%)	10 (13.7%)
Low risk	0 (0%)	5 (15.2%)	63 (86.3%)
	ADOS classification		
	Autism	PDD-NOS	DD
T-STAT risk category	(*n* = 54)	(*n* = 20)	(*n* = 62)
High risk	51 (94.4%)	15 (75%)	2 (3.2%)
Low risk	3 (5.6%)	5 (25%)	60 (96.8%)